

Przemysław Zawadzki, Uniwersytet Jagielloński

Zarys filozoficzno-etyczno- -społecznych implikacji rozwoju interfejsów mózg-komputer i mózg-mózg

Philosophical, Ethical and Social Implications
of Brain-to-Computer and Brain-to-Brain Interfaces

Streszczenie

Etyczno-społeczne implikacje płynące z badań nad interfejsami mózg-komputer i mózg-mózg stanowią dziś wyzwanie dla filozofów i etyków. Niosą ze sobą potencjał przekształcenia społeczeństwa w różnych obszarach jego funkcjonowania. Z jednej strony dają nowe możliwości terapeutyczne, perspektywy usprawnienia komunikacji międzyludzkiej lub nawet stworzenia nowych form międzygatunkowego porozumiewania się. Jednak z drugiej strony stanowią fundamentalne zagrożenie dla wolności myśli jednostki i poczucia jej autonomii. Autor prezentuje w artykule możliwości interfejsów poprzez przegląd związanych z nimi przełomowych eksperymentów. Zarysowane tło eksperymentalne i teoretyczne będzie stanowić podstawę głównego celu pracy, czyli refleksji nad kluczowymi filozoficznymi dylematami, jakie stawiają rozwój i szeroka aplikacja wspomnianej technologii.

Słowa kluczowe: interfejsy mózg-komputer, interfejsy mózg-mózg, etyka technologii, neuroprywatność, wieloagentowość.

Abstract

Socio-ethical implications that emerged from brain-computer and brain-to-brain interfaces research exemplify one of the greatest challenges for contemporary philosophers and ethics. This advanced technology may exert extensive influences on many areas of human life. Its therapeutic potential and improvement of interpersonal or even interspecies communication will be widely applied in the future. On the other hand, interfaces may pose a fundamental danger for human freedom and autonomy. In this article, the presentation of interfaces technology and their future potential will be provided. This outline, will lead to a major aim of the article, namely a reflection on the philosophical dilemma of a key importance.

Keywords: brain-computer interfaces, brain-to-brain interfaces, ethics of technology, neural privacy, multiagency.

Krótką historia interfejsów

Bezpośrednia komunikacja pomiędzy mózgiem a zewnętrznym urządzeniem okazała się faktem w latach 60. XX wieku, kiedy William Grey Walter przez przypadek stworzył działający interfejs (za: Cudo, Zabielska, Bałaj 2011: 95). Podczas jednej z operacji zaimplantował elektrody do obszaru kory motorycznej pacjenta, aby rejestrować jego aktywność neuronalną, podczas gdy badany miał przełączać slajdy na rzutniku za pomocą przycisku. Następnie dr Walter zmienił ustawienia systemu w taki sposób, aby projektor reagował, gdy zapis aktywności elektrycznej mózgu pokaże, że osoba chce nacisnąć przycisk. W tych okolicznościach badacz zauważył, że slajd zmieniał się, zanim pacjent zdążył wcisnąć przycisk.

Pięć lat po tym epizodycznym, nieopublikowanym wyniku Eberhard Fetz (1969), przeprowadzając eksperymenty na małpach, pokazał, że jeżeli są one nagradzane za generowanie odpowiednich wzorców aktywności neuronalnej, to przy zaimplementowanych elektrodach do obszaru kory motorycznej i odpowiednio zaprojektowanym układzie analogowym, który będzie transformował rejestrowane impulsy napięciowe na wychylenia igły woltomierza, są one zdolne w szybkim tempie nauczyć się rozmyślnie kontrolować urządzenia zewnętrzne świadomą modulacją poszczególnych neuronów.

Systematyczne badania poświęcone interfejsom mózg-komputer stały się rzeczywistością w latach 70. XX wieku za sprawą Jacques'a Vidala (Vidal 1973, 1977). W swoich eksperymentach Vidal,

do określenia kierunku spojrzenia osoby badanej, a tym samym przewidzenia kierunku, w jakim chce przesunąć kursor, wykorzystywał potencjały rejestrowane z czaszki w okolicy wzrokowej. Wyniki i refleksje nad tymi badaniami przełożyły się na pierwsze naukowe publikacje z dziedziny interfejsów (Cudo, Zabielska, Bałaj 2011: 7–8).

Jednak na pierwszą międzynarodową konferencję poświęconą tej tematyce trzeba było poczekać aż do 1999 roku. Tematy podjęte na konferencji dobrze obrazują główne tło problemowe, z jakim zmagają się badacze interfejsów aż do teraz, tj.: dopasowaniem interfejsów do potrzeb użytkowników, sposobami sterowania, potencjałem inwazyjnych i nieinwazyjnych metod, analizą sygnału rejestrowanego podczas aktywności neuronalnej, algorytmami służącymi do przetworzenia rejestrowanego sygnału na polecenia sterujące urządzeniem czy wymiarem praktycznym zastosowania technologii (Wolpaw i in. 2000).

Spotkanie w Rensselaerville Institute of Albany, przyczyniając się do integracji badaczy, wpłynęło na prędkość rozwoju w zakresie technologii interfejsów. Zgodnie z tym trendem w latach 2014–2015 udało się stworzyć interfejs pozwalający na bezpośrednią komunikację na linii mózg–mózg¹. Jednak zanim zostaną przedstawione szczegóły, warto najpierw ogólnie zarysować sposób działania interfejsów.

Ogólny zarys działania interfejsów

Interfejs mózg–komputer, najbardziej ogólnie, jest systemem, którego działanie opiera się na zamianie sygnału powiązanego z aktywnością mózgu na sygnał cyfrowy, który następnie zostaje wykorzystany do sterowania aplikacjami lub też elektronicznymi urządzeniami w czasie rzeczywistym. Zależnie od realizacji interfejsy mogą składać metaboliczną, magnetyczną lub elektryczną aktywność

¹ Bezpośrednia komunikacja oznacza tutaj komunikację z pominięciem jakichkolwiek modalności zmysłowych.

mózgu (Henle i in. 2012: 86). Zazwyczaj BCI transformuje elektrofizjologiczne sygnały z systemu nerwowego na przewidywane produkty tej aktywacji poprzez zakodowanie ich na wiadomości i komendy, które wpływają na realny świat. Ta charakteryzacja koresponduje z ujęciem interfejsów przez Wolpawa i in. (2002), którzy sądzą, że BCI jest urządzeniem zapewniającym mózgowi całkowicie nowy, gdyż niezależny od działania mięśni, kanał wpływania na świat zewnętrzny i komunikacji z innymi podmiotami.

W typowym projekcie BCI łączy się technologię obrazowania mózgu² z technologią uczenia maszynowego, aby automatycznie rozpoznawać wzory aktywności mózgu, przypisując je do konkretnych mentalnych zadań (Vleck i in. 2012). W większości BCI system jest wytrenowany do rozpoznawania tylko małego podzbioru czynności mentalnych³. Kiedy użytkownik wykonuje jedną z takich operacji, automatycznie zapoczątkowywana zostaje przypisana jej czynność⁴.

Postuluje się (Cudo, Zabielska, Bałaj 2011; Rouse, Shieber 2015), że każdy BCI składa się z czterech podstawowych komponentów:

1. Jako sygnał wejścia do BCI służy zarejestrowana aktywność mózgu osoby wykonującej daną mentalną czynność.
2. Za pomocą odpowiednich procedur, np. pomiaru amplitudy napięcia czy analizy widma, swoiste cechy są ekstrahowane z sygnału cyfrowego.
3. W celu zidentyfikowania mentalnej operacji wykonywanej przez użytkownika następuje dopasowanie zarejestrowanego sygnału do danych wzorców aktywności neuronalnej.
4. Końcowym etapem w systemach BCI jest dwutorowe działanie obejmujące sformułowanie polecenia sterującego urządzeniem, a także przekazanie informacji zwrotnej osobie używającej BCI.

W przypadku BTBI, w ostatnim ze wspomnianych wyżej punktów, działanie dotyczące polecenia sterującego obejmuje, zamiast urządzenia, mózg drugiej osoby (Perbal 2015).

² Zazwyczaj jest to elektroencefalografia (EEG).

³ Zazwyczaj jest to wyobrażanie sobie poruszania ręką czy nogą.

⁴ Jak np. poruszenie kursorem na ekranie monitora.

W zależności od stopnia ingerencji można wyróżnić inwazyjne i nieinwazyjne BCI i BTBI. Inwazyjne interfejsy są tworzone poprzez chirurgiczny zabieg wszczepienia sensorów do mózgu. Urządzenia te zapewniają najlepszy sygnał⁵, jednak stanowią ryzyko potencjalnych uszkodzeń mózgu. Ze względu na zagrożenia techniki te są znacznie ograniczane i nie stosuje się ich szeroko w celach medycznych (Tamburrini 2009). Zdecydowanie bezpieczniejsze i w związku z tym popularniejsze są nieinwazyjne interfejsy⁶, które nie wymagają ingerencji chirurgicznych. Używa się w nich zazwyczaj EEG w celu mierzenia elektrycznej aktywności powiązanej z funkcjonowaniem mózgu czy też funkcjonalnego rezonansu magnetycznego (fMRI) do obrazowania mózgu⁷ (Allison 2011). Przedstawione w artykule eksperymenty będą dotyczyły obydwóch metod, przy czym wszystkie zaprezentowane badania wykonywane na ludziach były przeprowadzane za pomocą nieinwazyjnych metod.

Wyniki badań na szczurach

Pierwszym badaniem ukazującym, że możliwy jest bezpośredni transfer informacji między mózgami, był przeprowadzony za pomocą inwazyjnych metod eksperyment Pais-Vieiry i in. (2013). Dokonali oni transferu informacji sensomotorycznych z mózgu jednego szczura do mózgu drugiego szczura. Procedura wyglądała następująco: pierwszy ze szczurów został poddany zadaniu, w którym miał do wyboru dwie dźwignie; jeśli wybrał tę, nad którą świeciło się światółko, zostawał nagradzany. Po pewnym czasie warunkowania instrumentalnego potrafił on z 96-procentową

⁵ Który jest nieosiągalny za pomocą technik nieinwazyjnych.

⁶ Stanowią około 80 proc. wszystkich interfejsów (Allison 2011).

⁷ Interfejsy wykorzystujące EEG są faworyzowane ze względu na ich funkcjonalność, przenośność i krótszy czas przygotowywania za ich pomocą procedury (Tamburrini 2009).

skutecznością wybierać dźwignie, za której przyciśnięcie został poddawany gratyfikacji. Celem badania było przesłanie informacji z mózgu wytrenowanego szczura do mózgu drugiego szczura, który znajdował się w analogicznie urządzonej klatce obok. Oczywiście szczury były pozbawione jakiejkolwiek możliwości zmysłowego kontaktu, co oznacza, że klatki były całkowicie oddzielone. W przypadku sukcesu oba szczury odbierały nagrodę. Architektura tego systemu w uproszczeniu wyglądała tak: szczur pierwszy, czyli agent kodujący informację, został wyposażony w elektrody rejestrujące aktywność jego podstawowej kory ruchowej. Aktywność ta zostawała zbierana, a następnie przesyłana do drugiego szczura, czyli agenta odkodowującego, wyposażonego w elektrody, dzięki którym można było stymulować konkretne obszary za pomocą odpowiedniego wzoru zebranej aktywności neuronalnej szczura pierwszego. Dzięki przekazanej informacji szczur odkodowujący potrafił w około 65 proc. wybrać prawidłową dźwignię, przy czym w niektórych sesjach wynik nie spadał poniżej 80 proc.⁸

Aby w pełni przetestować potencjał BTBI, ci sami badacze (Pais-Vieira i in. 2013) postanowili przesłać dane na większą odległość. W tym celu umieścili jednego szczura na uniwersytecie w IINN-ELS w Brazylii, a drugiego w laboratorium Uniwersytetu Duke'a w Stanach Zjednoczonych. Kodujący szczur, znajdujący się w IINN-ELS, wykonywał czynności analogiczne do poprzedniego eksperymentu. Następnie zebrana neuronalna aktywność została wysłana za pomocą Internetu do drugiego szczura (laboratorium Uniwersytetu Duke'a). Odpowiedni algorytm zamienił przetransferowany potencjał akcji we wzór stymulacji. Kiedy agent odkodowujący dokonał selekcji dźwigni, informacja

⁸ Głównym czynnikiem wpływającym na wynik odkodowującego szczura była jakość przestrzennej informacji pozyskanej od szczura kodującego. Im dokładniejsze było odróżnienie strony, po której należało nacisnąć odpowiednią dźwignię, i dłuższy czas trwania dostarczanych impulsów, tym wyższy wynik szczura odkodowującego.

zwrotna zostawała wysłana do klatki w IINN-ELS. Wyniki tego badania, oprócz ponownego potwierdzenia możliwości interfejsów w przesyłaniu informacji w czasie rzeczywistym z pominięciem jakichkolwiek zmysłów, ukazały, że odległość, w jakiej znajdują się agenci kognitywni, nie jest dla tej technologii ograniczeniem.

W tym miejscu warto wspomnieć o podobnym systemie testowanym przez Deadwylera i in. (2013). Nie był on oparty na BTBI, lecz obejmował podobny projekt, w którym występują nadawca i odbiorca informacji. Eksperyment zespołu Deadwylera zaprezentował niezwykle istotną cechę wydobywania wzoru aktywacji neuronów w trybie online z komórek znajdujących się w hipokampie⁹. Polegało to na pokazaniu, jak zmiana i torowanie funkcji hipokampa może być przeprowadzone poprzez bezpośrednie przekazanie informacji z tej struktury od szczura kodującego do drugiego szczura (odkodowującego). Zespół Deadwylera (2013) zaprojektował następującą procedurę badania: nadawca (pierwszy szczur) został wytrenowany w ściśle określonym zadaniu. Kiedy opanował on już określony wzór behawioralny, za pomocą skonstruowanego matematycznego modelu wyprowadzono wzór aktywności neuronalnej trenowanego szczura i wywołano go poprzez elektryczną stymulację w mózgu¹⁰ drugiego szczura, który nigdy nie był trenowany w wykonywaniu tego zadania. Okazało się, że nietrenowany szczur potrafił odtworzyć określoną sekwencję zadania, co może oznaczać, że przesłany został ślad pamięciowy zawierający „algorytm” jego wykonania.

⁹ Czyli obszarze, o którym od dawna było wiadomo, iż pełni u ssaków kluczowe funkcje, jeśli chodzi o konsolidację informacji z pamięci krótkotrwałej do pamięci długotrwałej, a także jeśli chodzi o pamięć deklaratywną i orientację przestrzenną. Ostatnio zidentyfikowano dokładny mechanizm, który demonstrował, że dekodowanie informacji podczas wykonywania zadania związanego z pamięcią krótkotrwałą było zapewniane przez określony wzór aktywności generowanej przez grupę neuronów znajdujących się właśnie w hipokampie (za: Deadwyler 2013).

¹⁰ Dokładniej: w hipokampie.

Transfer informacji pomiędzy ludźmi

Zespół Grau (2014) dokonał pierwszego eksperymentu, który wykazał świadomą transmisję informacji za pomocą interfejsu mózg-komputer-mózg między ludźmi. Interfejs *Brain-to-Brain* został skonstruowany za pomocą zainstalowania badanym znajdującym się w Indiach BCI wykorzystującego EEG, a we Francji CBI opartego na TMS (przeznaczszkowa stymulacja magnetyczna). Przepływ informacji, analogicznie jak w przypadku szczurów, umożliwiło połączenie z Internetem. Pierwszy z uczestników eksperymentu, tj. nadawca, za pomocą poruszenia ręką lub stopą kodował informację o lub 1. Następnie jego aktywność neuronalną zbierano za pomocą EEG i transmitowano do odbiorcy. Odbiorca informacji został poddany dwufazowej stymulacji odpowiednich obszarów kory potylicznej. W zależności od wartości zakodowanej informacji w odpowiedni sposób skierowywano pole TMS. Wartość 1, czyli kierunek aktywny, wywoływała doświadczenie percepcji światła, dla której zakodowano słowo „Hola”. Wartość 0 wpływała na prostopadły (ortogonalny) kierunek stymulacji, co nie miało przełożenia na doświadczenie odbioru światła. Ten sposób stymulacji był zakodowanym słowem „Ciao”. Następnie odbiorca zdawał słowny raport o wystąpieniu u niego doświadczenia światła bądź o jego braku. Można to uznać za załazek nowej interkontynentalnej formy komunikacji, w której za pomocą ruchu ręką można przesłać powitalną wiadomość bezpośrednio do mózgu.

Podobne badanie z wykorzystaniem tych samych technologii wykonał Rao i in. (2014). Jednak zamiast powitania zakodowanego w doświadczeniu światła badani mieli za zadanie współpracować w odpowiedzi na pytania prezentowane na ekranie. Interfejs BTB wykrywał obrazowanie aktywności kory motorycznej przedstawionej przez sygnał EEG zebrany od pierwszego z badanych (nadającego informację agenta) i transmitował ją przez Internet do kory motorycznej drugiej osoby (odbierającego informację agenta). To pozwalało nadającemu informację wywołać pożądane reakcje u odbierającego informację (przyciśnięcie panelu dotykowego).

Okazało się, że odbierający informację badany odpowiadał na pytania z 72-procentową skutecznością, w porównaniu z badaniem kontrolnym, tj. bez przekazywania informacji¹¹.

Eksperymenty te demonstrują możliwości BTBI do komunikacji nowej generacji, a nawet do kontroli za pomocą samej aktywności mózgu zachowania drugiej osoby. Tak jak miało to miejsce w przypadku raportu z badania Sakurada i in. (2013), w którym EEG było wykorzystane do kontroli interfejsu mózg–maszyna, zainstalowanego w kombinezonie pierwszego użytkownika, a kierowane w celu łapania i rzucania piłeczki przez drugiego z badanych. Zdaniem Rao i in. (2014) powyższe eksperymenty to pierwszy krok na drodze ku przełamaniu tradycyjnego, opartego na języku porozumiewania się, co otwiera drogę nowym kierunkom badawczym, włączając w to na przykład tak ciekawe formy przekazywania informacji, jak transmisja emocji i uczuć. Już teraz bada się technologie sprzyjające jej skonstruowaniu w przyszłości (Bajaj, Pachor 2015) czy też, co wydaje się najbardziej skomplikowane, transmisję treści konceptualnych.

Interfejsy wieloagentowe i intergatunkowe

Niedawno dzięki Ramakrishnanowi i in. (2015) udało się dokonać przełomu w badaniach nad interfejsami wieloagentowymi. Jak dowiadujemy się z poprzednich badań, tradycyjne interfejsy wydobywały ruchowe decyzje bądź rozkazy podmiotu tylko z jednego mózgu: czy to do kontroli urzędzeń, czy do kontroli innej osoby. Natomiast we wspomnianym badaniu okazało się, że jest możliwe zbudowanie internetu mózgów (brainnet). Wykorzystuje on wielkoskalową aktywność mózgu nawet trzech prymatów, aby

¹¹ To wynik znaczący, gdyż w tym drugim skuteczność oscylowała w granicach 18 proc. Brak 100-procentowej skuteczności mógł być spowodowany niedoskonałością aparatury i/lub nieprawidłową interpretacją sygnałów przez mózg odbiorcy.

wciągnąć je we wspólne osiąganie wyznaczonego celu, jakim jest np. ruszanie ręką robota. Okazuje się, że prymaty osiągają zdecydowanie lepsze wyniki, kiedy wykonują to zadanie razem niż osobno.

Jest to zgodne z innymi badaniami (Pais-Vieira i in. 2015). Między innymi okazało się, że połączenie czterech mózgów szczurów sprawia, że są one mądrzejsze, niż kiedy pozostają odosobnione. Szczury wykonywały zadania z klasyfikacji obiektów, rozpoznawania obrazów, przechowywania i przetwarzania informacji ruchowych czy nawet przewidywania pogody. Przykład pogody jest oczywiście pewnym zabiegiem mającym na celu popularyzację technologii. Polega on na tym, że zwiększona lub zmniejszona liczba impulsów dostarczana do mózgu reprezentuje informacje o ciśnieniu i temperaturze, i to na podstawie tak zakodowanych informacji szczury, które są poddane warunkowaniu, prognozują, czy będzie padać.

Z perspektywy namysłu nad etycznymi implikacjami interfejsów warto wspomnieć o pomysłe zrealizowanym przez zespół profesora Yoo (2013) z Uniwersytetu Harvarda. Badania na medycznym wydziale tej uczelni doprowadziły do skonstruowania interfejsu intergatunkowego. Profesor Yoo użył w tym celu EEG do rejestru aktywności neuronalnej osób, którym prezentowano bodziec wizualny. Kiedy badany skupiał się intencjonalnie, wywoływało to silny sygnał rozpoczynający stymulację odpowiednich ośrodków mózgu szczura, co powodowało poruszanie ogonem. Profesor Yoo przedstawia także perspektywę odwrócenia sytuacji, w której to szczur będzie poruszał kończynami człowieka. Badania takie mogą mieć wymiar pragmatyczny, mogą prowadzić do zwiększenia zakresu odbieranych informacji zmysłowych, np. poprzez interfejs z mózgiem psa byłyby to informacje węchowe. W tym kontekście rodzi się refleksja, aby interfejsy, w których ludzie są sprzężeni z psami, wykorzystywać na przykład w akcjach ratunkowych podczas odnajdywania zaginionych (Trimper i in. 2014).

Zastosowania interfejsów

Ludzkość może wiele zyskać dzięki interfejsom. Coraz bardziej rozwijają się technologie wspomagające komunikację ludzi z tzw. syndromem zamknięcia, co dla cierpiących na tę chorobę może być jedyną szansą na bardziej złożoną formę komunikacji (Nijboer, Broermann 2010). Interfejsy wspomagają także kierowanie neuroprotezami (Müller-Putz i in. 2010). W samej Europie na uszkodzenia rdzenia kręgowego cierpi około 300 tysięcy ludzi, a co roku dochodzi do tego nowych 11 tysięcy urazów (Peckham 2001; za: Müller-Putz 2010). To sprawia, że rozpowszechnienie omawianych technologii w celach terapeutycznych jest bardzo istotnym społecznym celem.

Także w edukacji technologie interfejsów mogą okazać się rewolucyjne, i to w różnych jej obszarach. Po pierwsze, obiecujące są wyniki badań neurofeedbacku (Neuper, Pfurtscheller 2010). Po drugie, przy rozwinięciu technologii zaprezentowanej w badaniu Deadwylera i in. (2013), podobnie jak w przypadku szczurów, u których wywołano wzory behawioralne za pomocą przekazanej informacji, być może zintegrowane mózgi nauczyciela i ucznia pozwolą na szybsze nabycie umiejętności przez tego drugiego. Po trzecie, można spekulować, że zintegrowanie mózgu dwóch uczonych, analogicznie do „internetu mózgów” (Pais-Vieira in. 2015) w przypadku zwierząt, może prowadzić do lepszych rezultatów badawczych.

BCI jest już wykorzystywane także w branży rozrywkowej. Z jego użyciem można w bardzo ograniczonym zakresie próbować swoich sił w popularnej grze MMO World of Warcraft, wykonywać komendy przypisane przyciskom klawiatury bez poruszania rękami, tj. tylko za pomocą myśli, przeprowadzać trening kognitywny (pracować nad polepszeniem uwagi i pamięci) czy obsługiwać specjalne aplikacje zainstalowane na smartfonie. Co ciekawe, deweloperzy aplikacji mają dostęp: do wytwarzanych w ten sposób danych (Martinovic i in. 2012).

Dylematy natury moralnej

Technologie interfejsów oprócz niewątpliwych korzyści budzą także mnóstwo dylematów moralnych. Dodatkowo atmosferę podgrzewają publikacje prasowe z obszaru marketingu, jak ta, która pojawiła się w 2008 roku. The Nielsen Company, która nabyła firmę NeuroFocus specjalizującą się w neuronauce, rozwinęła za jej pomocą urządzenie o nazwie Mynd. Naukowcy i przedsiębiorcy związani z tym projektem sądzą, że dzięki analizie danych internetowych będą w stanie zrozumieć i wykorzystać do celów komercyjnych głębokie i podświadome reakcje konsumentów, które pojawiają się w odpowiedzi na konkretne treści. Zapowiadają oni, że pozwoli to na budowanie głębokiej i trwałej relacji z klientem, co zrewolucjonizuje rynek sprzedaży¹². W tym miejscu warto zastanowić się nad szerszym spektrum zagrożeń płynących z użycia technologii interfejsów (Glannon 2014; Tamburrini 2009; Hildt 2011, 2015). Najważniejsze z problemów stojących przed społeczeństwem to refleksja nad neuroprywatnością, poczuciem tożsamości (Trimper i in. 2014), a także nad wspomaganiem ludzkich procesów poznawczych (Bostrom, Sanders 2009).

Aby przedstawić zagrożenia, które może wywołać pierwsza z wymienionych kategorii, neuroprywatność, nie trzeba wybiegać daleko w przyszłość. Można odwołać się do badania wykonanego już przed kilkoma laty (Martinovic i in. 2012). Oparto je na wykrywaniu fali mózgowej P300, składającej się z dwóch subkomponentów, nazywanych falą P3a i P3b. Obie są związane ze złożonymi kognitywnymi procesami, jak rozpoznawanie i kategoryzacja bodźców płynących ze środowiska. Dzięki znajomości własności tych fal badacze sądzili, że gdy ludzie rozpoznają coś, co ma dla nich znaczenie (np. znajomego czy zdjęcie), to pojawi się fala P300. Dzięki odpowiedniej prezentacji bodźców faktycznie okazało się możliwe, aby w ten sposób wydobyć od badanych prywatne

¹² <http://www.nielsen.com/us/en/solutions/capabilities/consumer-neuroscience.html> (dostęp: 24.05.2016).

informacje, takie jak: kod PIN, kod karty debetowej, nazwę banku, w którym badani posiadają rachunek, miesiąc urodzenia, znajomość osób czy miejsce zamieszkania. Do zebrania aktywności neuronalnej (która wystąpiła, gdy badanym wyświetlano na ekranie różne obrazy w celu zdobycia poufnych informacji) użyto EEG. Eksperyment rozpoczął się od krótkiego treningu. W pierwszym przypadku użytkownik po części aktywnie kalibruje¹³ BCI, w drugim przypadku jest pasywny, co oznacza, że nie bierze aktywnego udziału w treningu systemu. W drugim z przypadków (pasywnym) eksperymentatorzy wykorzystywali specjalne klasyfikatory¹⁴. Następnie badanym prezentowano odpowiednie bodźce, w zależności od tego, które informacje chciano od nich uzyskać. Skuteczność w odgadywaniu prawidłowych odpowiedzi w przypadku eksperymentów dotyczących kodu PIN, kart debetowych czy znajomych wynosiła 20 proc., w przypadku lokalizacji zamieszkania i nazwy banku 30 proc., a miesiąca urodzenia aż 60 proc.¹⁵. Te stosunkowo niskie wyniki były spowodowane specyfiką procedury. W podobnym badaniu, nazywanym Guilty-Knowledge Test (GKT), zastosowano nieco odmienną procedurę. Okazało się, że skuteczność detekcji pożądaných informacji osiągnęła 86 proc. Różnice między badaniami dotyczyły głównie podejścia do modelu ataku. W GKT osoba przesłuchująca ma pełną kontrolę nad użytkownikiem BCI, może rozmieścić elektrody o wysokiej rozdzielczości i zmusić

¹³ Kalibracja polegała na prezentacji użytkownikom losowej sekwencji numerów od 0 do 9. Ponadto zostali proszeni o zliczenie powtórzeń wybranego numeru (targetu). Każdy numer był wyświetlany 16 razy, czas trwania bodźca wynosił 250 ms, a przerwa losowo od 250 ms do 375 ms. Na końcu eksperymentu badani byli proszeni o podanie swoich zliczeń w celu sprawdzenia poprawności wyniku.

¹⁴ Szerzej przedstawione w artykule (Martinovic i in. 2012).

¹⁵ Wyniki różniły się między różnymi klasyfikatorami, a także w zależności od tego, czy badany trenował system aktywnie czy pasywnie. Tak wysoki wynik w przypadku identyfikacji miejsca zamieszkania był konsekwencją specyficznie scharakteryzowanego celu zadania, w którym badany był proszony o zliczanie powtórzeń celu (targetu).

badanego do współpracy (za: Martinovic i in. 2012). W badaniu zespołu Martinovica (2012) atakujący miał do dyspozycji tani sprzęt do grania¹⁶ wybrany przez użytkownika¹⁷.

Wiemy już także z przeprowadzonych lata temu eksperymentów, że można wykorzystać funkcjonalny rezonans magnetyczny (fMRI) do rekonstruowania oglądanych przez badanych filmików (Nishimoto i in. 2011), dekodowania kategorii obiektów wyświetlanych uczestnikom badań (za: Trimper, Wolpe i Rommelfanger 2014) czy wykrywania tego, czy kłamią (za: Trimper, Wolpe i Rommelfanger 2014). Technologie interfejsów budzą jednak obawy innej natury (Trimper, Wolpe i Rommelfanger 2014). Mogą bowiem prowadzić nie tylko do dekodowania informacji z mózgu, lecz także do stymulacji mózgu w sposób, w którym podmiot nie będzie mógł się oprzeć bodźcowi, jak w przypadku doświadczeń przeprowadzanych na myszach przez zespół Ramirez (Ramirez i in. 2013), gdzie wszczepiono myszy fałszywe wspomnienia. W tej sytuacji można sobie łatwo wyobrazić zagrożenia, gdy użytkownik podłącza się do Internetu za pomocą BCI. Daje to pole manewru neurohackerom.

Obawy związane z drugą kategorią, czyli możliwością utraty poczucia tożsamości, a także z kategorią trzecią, która odnosi się do augmentyki, opierają się na biokonserwatywnym przeświadczeniu, że BCI, które nie jest związane ściśle z terapią, może doprowadzić do utraty tego, co ludzkie, lub przekroczeniu pewnej granicy naturalności, za którą ludzie niepewni konsekwencji swoich działań mogą stanowić niebezpieczeństwo dla siebie samych.

Implikacje rozwoju interfejsów to tak naprawdę stare wyzwania filozofii w nowej szacie. Jeśli przyjmiemy często akceptowane

¹⁶ Było to popularne urządzenie Emotiv EPOC BCI, dostępne na rynku od 2009 roku.

¹⁷ Co więcej, istnieją sposoby niewykorzystywane w badaniach, jak globalne statystyki, które mogą zdecydowanie poprawić szanse na sukces ataku (np. dystrybucja klientów banków w populacji jest znana w przybliżeniu). Ponadto można dysponować wiedzą zdobytą wcześniej, która zawęży obszar poszukiwań.

współcześnie stanowisko filozoficzne, że tożsamość osobową można sprowadzić do ciągłości życia psychicznego¹⁸, to co się stanie, kiedy zsynchronizujemy swój mózg z mózgiem lub mózgami innych osób? Czy dojdzie wtedy do połączenia we wspólną, wieloosobową tożsamość? A jeśli tak, to czy wieloosobowa tożsamość będzie istniała tylko przez okres podłączenia do takiego interfejsu czy zaistnieje bardziej stała relacja pomiędzy podmiotami biorącymi udział w eksperymencie? W kontekście wspomnianej teorii tożsamości niezwykle istotna staje się możliwość świadomego monitorowania przeżyć, gdyż tworzą one aktualne i przeszłe „ja”, czyli podmiot trwania psychicznego. W tym względzie wieloagentowy interfejs może prowadzić do efektów podobnych jak dla zjawiska klinicznego znanego jako „myśli nasłane”, a więc stwarzać przekonania¹⁹ o własnej bierności, co jest właśnie zaburzeniem zdolności monitorowania własnych doświadczeń (Bremer 2012). Jedna z cierpiących na to zaburzenie osób tak opisuje swoje przeżycia:

¹⁸ Parfit (2001) wyróżnia dwie koncepcje tożsamości osoby: prostą i złożoną. Koncepcja prosta odpowiada potocznym intuicjom, ponieważ zakłada świadomość tożsamości i istnienie faktów transcendentnych, czyli niezależnych od empirycznej ciągłości. Tożsamość stanowi w tym ujęciu coś istotowo stałego, co tworzy każdego z nas. W ujęciu złożonym świadomość tożsamości jest natomiast mniej podstawowa i często eliminowana. Zastępuje się ją pojęciami trwania podmiotu pod względem fizycznym lub psychicznym. Współcześnie, między innymi po debatach nad możliwymi konsekwencjami teleportacji czy też przeniesienia świadomości na inny nośnik, trwanie fizyczne często nie jest już postrzegane jako warunek zachowania tożsamości. W związku z tym istnieje tendencja do funkcjonalistycznego podejścia do fizycznego podłoża (czy to krzem, czy białko ważne, żeby zaistniała psychika). Stąd przywołane w tekście stanowisko, iż tożsamość można sprowadzić do ciągłości życia psychicznego, które w uproszczeniu można sprowadzić do ciągłości inklinacji, celów, wspomnień (Bremer 2012).

¹⁹ W przypadku „nasłanych myśli” są to fałszywe przekonania (urojenia), natomiast w przypadku wieloagentowego interfejsu będą to prawdopodobnie przekonania prawdziwe.

Patrzę przez okno i myślę, że ogród wygląda miło, a trawa wygląda ładnie, lecz myśli Eamona Adrewsa wchodzą do mojego umysłu [...] używa on mojego umysłu jak ekranu i wyświetla na nim swoje myśli (za: Bremer 2012).

Być może agenci podłączeni do zaawansowanego interfejsu wieloagentowego będą raportowali podobne doświadczenie fenomenologiczne. Co stanie się wtedy z kategorią odpowiedzialności osobowej, na której w dużej mierze była oparta kultura europejska? Kto będzie odpowiedzialny za czyny podmiotów podłączonych do interfejsu wieloosobowego? Co z potencjalnymi uszkodzeniami mózgów uczestników badania? Czy etyczny ciężar takich badań spada całkowicie na eksperymentatora (Grübler 2011; O’Brolchain, Gordijn 2014)?

Oprócz debaty naukowo-etycznej mającej na celu udzielenie odpowiedzi na te pytania bardzo ważną kwestią są odpowiednie regulacje prawne. Istotne jest rozważenie zakresu²⁰, w którym powinno być dopuszczone stosowanie interfejsów w przesłuchiowaniu obywateli przez władze. Następnym ważnym krokiem jest wywołanie międzynarodowej debaty dotyczącej konkretnych regulacji wykorzystywania interfejsów podczas wojny. Problem jest o tyle aktualny, że prowadzone są testy w zakresie przydatności interfejsów w tym zakresie w Amerykańskiej Agencji Zaawansowanych Projektów Badawczych w Obszarze Obronności (DARPA). Może to w niedługim czasie doprowadzić do użycia tych technologii w praktyce, aby bezpośrednio przekazywać informacje walczącym żołnierzom. Jeśli technologia ta zostanie wprowadzona, może mieć to pewne pozytywne skutki: precyzyjne instrukcje od żołnierza obserwującego front z filmu w czasie rzeczywistym mogą wpłynąć na zmniejszenie liczby przypadkowych zabójstw ludności cywilnej.

²⁰ I czy w ogóle powinien istnieć jakikolwiek dopuszczalny zakres. Z pewnością uzasadnione wydaje się pozyskanie w ten sposób informacji, jeśli mogą one pomóc uratować jednostkę bądź grupę osób (np. podczas zamachu terrorystycznego). Inne zastosowania będą budziły kontrowersje.

Jednak stwarza to także okazję hackerom chcącym wysłać komunikaty, które pokierują daną jednostką według ich woli (Trimper i in. 2014).

Ze względu na mnogość aplikacji interfejsów problemy z nimi związane są – oprócz modyfikacji genetycznych, rozwoju sztucznej inteligencji i refleksji nad wykorzystywaniem energii nuklearnej – najważniejszymi wyzwaniem początku XXI wieku, gdyż konsekwencje działania tych technologii są globalne. Z tego powodu debata nad interfejsami powinna zataczać szerokie kręgi i odbić się głośnym echem, aby publiczność niezwiązana na co dzień z neuronaukami mogła się konceptualnie przygotować na nadejście zarysowanych w artykule nowych form wpływu, jak możliwość bezpośredniego wpływu na zachowanie (Sakurada i in. 2012), manipulacji pamięcią (Ramirez i in. 2013) czy „kradzieży myśli” (Martinovic i in. 2012). Równie ważne jest wskazanie i konsekwentny rozwój obszarów, w których technologie interfejsów mogą dać społeczeństwu bardzo wiele pozytywnych i wartościowych rezultatów, jak na płaszczyźnie terapeutycznej (Nijboer, Broermann 2010; Müller-Putz i in. 2010), komunikacyjnej (Rao i in. 2014) czy edukacyjnej (Deadwyler i in. 2013; Pais-Vieira i in. 2015). Jak pokazuje historia, rozwoju technologii nie można zatrzymać w dowolnym momencie – dlatego działania popularyzacyjne muszą być przemyślane. Oznacza to, że wraz z propagowaniem nowych wyników technologicznych należy prowadzić także refleksję nad ich etycznymi i społecznymi implikacjami, gdyż to jedyny sposób na ograniczenie zgubnych wpływów i zintensyfikowanie korzyści płynących z rozwoju technologicznego.

Bibliografia

- Allison, B. (2011). *Trends in BCI Research: Progress Today, Backlash Tomorrow*. XRDS: Crossroads The ACM Magazine for Students.
- Bostrom, N.; Sandberg, A. (2009). *Cognitive Enhancement: Methods, Ethics, Regulatory Challenges*. Sci. Eng. Ethics, 15.

- Bajaj, V.; Pachor, R. (2015). *Detection of Human Emotions Using Features Based on the Multiwavelet Transform of EEG Signals*, [w:] A.E. Hassanien, A.T. Hazar (2015). *Brain-Computer Interfaces Current Trends and Applications*. Intelligent Systems Reference Library, 74. New York: Springer.
- Bremer, J. (2012). *Jaźń i tożsamość osobowa*, [w:] M. Miłkowski, R. Poczobut (red.), *Przewodnik po filozofii umysłu*, Kraków: WAM.
- Cudo, A.; Zabielska, E.; Bałaj, B. (2011). *Wprowadzenie w zagadnienie interfejsów mózg-komputer*, [w:] Studia z Psychologii w KUL. Wydawnictwo KUL. http://www.kul.pl/files/55/Stud_psych_18-2012_Cudo.pdf.
- Deadwyler, S.A.; Berger, T.W.; Sweatt, A.J. i in. (2013). *Donor/Recipient Enhancement of Memory in Rat Hippocampus*. *Front. Syst. Neurosci.*, 7.
- Fetz, E.E. (1969). *Operant Conditioning of Cortical Unit Activity*. *Science*, 163.
- Glannon, W. (2014). *Ethical Issues With Brain-Computer Interfaces*. *Front. Syst. Neurosci.*, 8.
- Grübler, G. (2011). *Beyond the Responsibility Gap. Discussion Note on Responsibility and Liability in the Use of Brain-Computer Interfaces*. *AI Soc.*, 26.
- Grau, C.; Ginhoux, R.; Riera, A. i in. (2014). *Conscious Brain-to-Brain Communication in Humans Using Non-Invasive Technologies*. *PLoS ONE*, 9.
- Henle, C.; Schuettler, M.; Rickert, J.; Stieglitz T. (2012). *Towards Electrocorticographic Electrodes for Chronic Use in BCI Applications*, [w:] B.Z. Allison, S. Dunne, R. Leeb i in. (2012). *Towards Practical Brain-Computer Interfaces. Bridging the Gap from Research to Real-World Applications*, Heidelberg: Springer.
- Hildt, E. (2011). *Brain-Computer Interaction and Medical access to the Brain: Individual, Social and Ethical Implications*. *Stud. Ethics Law Technol.*, 4.
- Hildt, E. (2015). *What Will This Do To Me and My Brain? Ethical Issues in Brain-to-Brain Interfacing*. *Front. Syst. Neurosci.*, 9.
- Martinovic, I., Davies, D., Frank, M. i in. (2012). *On the Feasibility of Side-Channel Attacks With Brain-Computer Interfaces*, *Proceedings of the 21st USENIX conference on Security symposium*, August 2012, Bellevue, WA.
- Müller-Putz, G.R.; Scherer, R.; Pfurtscheller, G.; Rupp, R. (2010). *Non Invasive BCIs for Neuroprostheses Control of the Paralyzed Hand*, [w:] B. Graimann, B. Allison, G. Pfurtscheller (eds.). *Brain-Computer Interfaces. Revolutionizing Human-Computer Interaction*. Berlin: Springer.

- Neuper, C.; Pfurtscheller, G. (2010). *Neurofeedback Training for BCI Control*, [w:] B. Graimann, B. Allison, G. Pfurtscheller (eds.) *Brain-Computer Interfaces. Revolutionizing Human-Computer Interaction*. Berlin: Springer.
- Nijboer, F.; Broermann, U. (2010). *Brain-Computer Interfaces for communication and control in locked in patients*, [w:] B. Graimann, B.Z. Allison, G. Pfurtscheller (eds.), *Brain Computer Interfaces Revolutionizing Human Computer Interaction*. New York: Springer Publishing.
- Nishimoto, S.; Vu, A.T.; Naselaris, T. i in. (2011). *Reconstructing Visual Experiences From Brain Activity Evoked by Natural Movies*. *Curr. Biol.*, 19.
- O’Brolchain, F.; Gordijn, B. (2014). *Brain-Computer Interfaces and User Responsibility*, [w:] G. Grübler, E. Hildt (eds.). *Brain-Computer Interfaces in their Ethical, Social and Cultural Contexts*. Dordrecht: Springer.
- Pais-Vieira, M.; Lebedev, M.; Kunicki, C. i in. (2013). *A Brain-to-Brain Interface for Real-Time Sharing of Sensorimotor Information*. *Scientific Reports*, 3.
- Pais-Vieira, M.; Chiuffa, G.; Lebedev, M.A. i in. (2015). *Building an Organic Computing Device with Multiple Interconnected Brains*. *Sci. Rep.*
- Parfit, D. (2001). *Uwagi na temat „Doniosłości tożsamości”*, przeł. R. Wiercok, [w:] J. Górnicka-Kalinowska (red.), *Filozofia podmiotu*, Warszawa: Aletheia.
- Perbal, B. (2015). *Knock Once for Yes, Twice for No*. *J Cell Commun Signal.*, 9(1).
- Ramakrishnan, A.; Ifft, P.J.; Pais-Vieira, M. i in. (2015). *Computing Arm Movements with a Monkey Brainet*. *Scientific Reports*, 5, Article number: 10767.
- Ramirez, S.; Liu, X.; Lin, P.A. i in. (2013). *Creating a False Memory in the Hippocampus*. *Science*, 341.
- Rao, R.P.; Stocco, A.; Bryan, M. i in. (2014). *A Direct Brain-to-Brain Interface in Humans*. *PLOS ONE*, 9(11).
- Rouse, A.; Schieber, M. (2015). *Advancing Brain-Machine Interfaces: Moving Beyond Linear State Space Models*. *Front. Syst. Neurosci.*, 28 July 2015, <http://dx.doi.org/10.3389/fnsys.2015.00108>.
- Sakurada, T.; Kawase, T.; Takano, K. i in. (2013). *Abmi-Based Occupational Therapy Assist Suit: A Synchronous Control by Ssvep*. *Front. Neurosci.*, 7.
- Tamburrini, G. (2009). *Brain to Computer Communication: Ethical Perspectives on Interaction Models*. *Neuroethics*, 2.

- Trimper, J.B.; Wolpe, P.R.; Rommelfanger, K.S. (2014). *When "I" Becomes "We": Ethical Implications of Emerging Brain-to-Brain Interfacing Technologies*. Front. Neuroeng., 7.
- Vidal, J.J. (1973). *Toward Direct Brain-Computer Communication*. Annual Reviews in Biophysics and Bioengineering, 2.
- Vidal, J.J. (1977). *Real-Time Detection of Brain Events in Eeg*. Proceedings of the IEEE, 65(5).
- Vlek, R.J.; Steines, D.; Szibbo, D. i in. (2012). *Ethical Issues in Brain Computer Interface Research, Development, and Dissemination*. J. Neurol. Phys. Ther., 36.
- Wolpaw, J.R.; Birbaumer, N.; Heetderks, W.J. i in. (2000). *Brain Computer Interface Technology: A Review of the First International Meeting*. IEEE Transactions on Rehabilitation Engineering, 8(2).
- Wolpaw, J.R.; Birbaumer, N.; McFarland, D.J. i in. (2002). *Brain-Computer Interfaces for Communication and Control*. Clinical Neurophysiology, 113.
- Yoo, S.S.; Kim, H.; Filandrianos, E. i in. (2013). *Non-Invasive Brain-to-Brain Interface (Bbi): Establishing Functional Links between two brains*. PLoS ONE, 8.